Biochimica et Biophysica Acta 1797 (2010) 406–413

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbabio

De novo design of a non-natural fold for an iron–sulfur protein: Alpha-helical coiled-coil with a four-iron four-sulfur cluster binding site in its central core

Joanna Grzyb [a,1], Fei Xu [b], Lev Weiner [c], Eduard J. Reijerse [d], Wolfgang Lubitz [d], Vikas Nanda [b], Dror Noy [a,*]

[a] Plant Sciences Department, Weizmann Institute of Science, Rehovot, Israel
[b] Biochemistry Department, Robert Wood Johnson Medical School, Univ. of Medicine and Dentistry of New Jersey, Piscataway, NJ, USA
[c] Chemical Research Support Department, Weizmann Institute of Science, Rehovot, Israel
[d] Max-Planck-Institut für Bioanorganische Chemie, Mülheim an der Ruhr, Germany

ARTICLE INFO

Article history:
Received 29 September 2009
Received in revised form 11 December 2009
Accepted 16 December 2009
Available online 24 December 2009

Keywords:
Iron sulfur cluster
Protein de novo design
Redox enzyme
Coiled-coil
Four-helix bundle
EPR spectroscopy

ABSTRACT

Using a 'metal-first' approach, we computationally designed, prepared, and characterized a four-iron four-sulfur (Fe_4S_4) cluster protein with a non-natural α-helical coiled-coil fold. The novelty of this fold lies in the placement of a Fe_4S_4 cluster within the hydrophobic core of a four-helix bundle, making it unique among previous iron–sulfur (FeS) protein designs, and different from known natural FeS proteins. The apoprotein, recombinantly expressed and purified from *E. coli*, readily self-assembles with Fe_4S_4 clusters *in vitro*. UV–Vis absorption and CD spectroscopy, elemental analysis, gel filtration, and analytical ultracentrifugation confirm that the protein is folded and assembled as designed, namely, α-helical coiled-coil binding a single Fe_4S_4 cluster. Dithionite-reduced holoprotein samples have characteristic rhombic EPR spectra, typical of low-potential, $[Fe_4S_4]^+$ ($S=1/2$), with g values of $g_{z,y}=(1.970, 1.975)$, and $g_x=2.053$. The temperature, and power dependence of the signal intensity were also characteristic of $[Fe_4S_4]^+$ clusters with very efficient spin relaxation, but almost without any interaction between adjacent clusters. The new design is very promising although optimization is required, particularly for preventing aggregation, and adding second shell interactions to stabilize the reduced state. Its main advantage is its extendibility into a multi-FeS cluster protein by simply duplicating and translating the binding site along the coiled-coil axis. This opens new possibilities for designing protein-embedded redox chains that may be used as "wires" for coupling any given set of redox enzymes.

1. Introduction

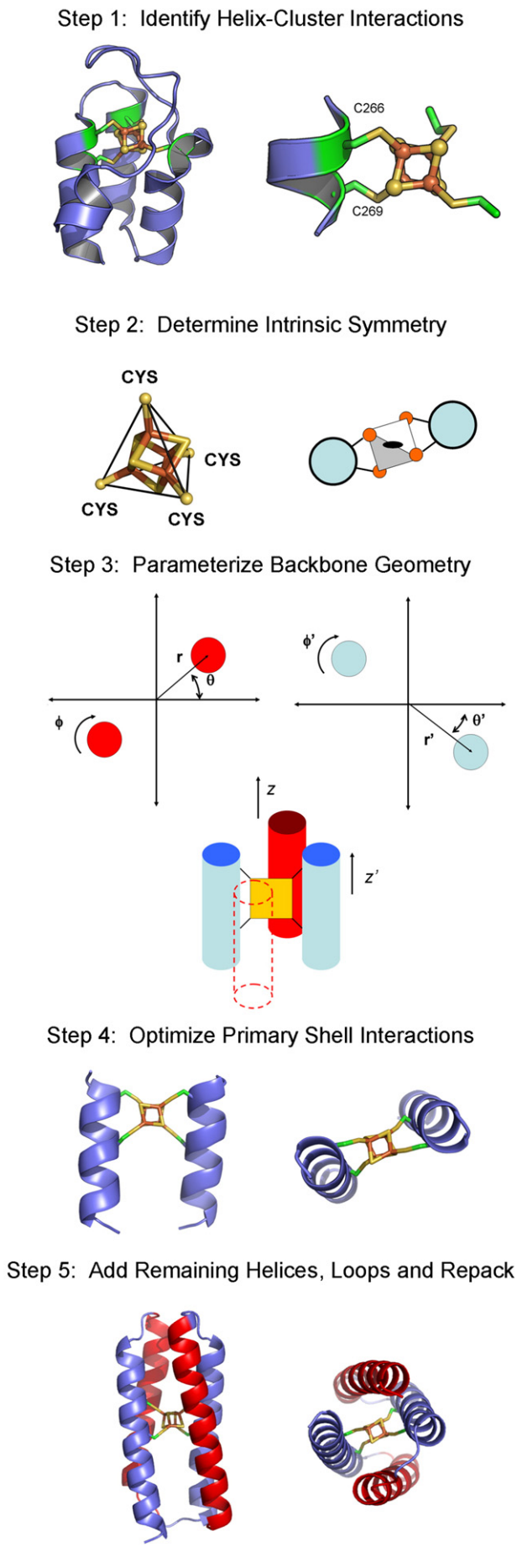
Iron–sulfur (FeS) proteins are ubiquitous redox proteins in all living organisms. Their structural simplicity, presence in simple anoxygenic bacteria, and usually highly reducing potentials and high oxygen sensitivity, suggest that they are evolutionary remains of the era before earth's atmosphere became oxygenic [1]. Although in present-day organisms they perform a variety of catalytic functions at a broader range of potentials, they are still used as carriers of electrons at highly reducing potentials; for example, in photosystem I (PSI) [2], complex I [3], and hydrogenases [4–6]. The growing interest in biological and photobiological hydrogen production as a clean renewable alternative for fossil fuels [7,8] is prompting greater interest in FeS proteins. Particularly, redox chains of four-iron four-sulfur clusters (Fe_4S_4), are receiving much attention as they facilitate long-range electron transfer to and from the catalytic centers of hydrogenases. This functionality is critical to any practical application that requires interfacing hydrogenases either to electrodes [9] or photoelectron sources [10,11].

In this context, de novo designed proteins are especially appealing for constructing electron-transfer relays between biocatalytic centers because they can be coupled to natural catalytic centers via native-like protein–protein interfaces [12]. With notable exceptions [13], most de novo designs are simple, robust folds, such as α-helical bundles. This makes multi-center FeS proteins especially difficult targets for de novo design because they are naturally found in dedicated loop regions with complicated folds. Most of the FeS-coordinating proteins are rich in beta structures [14]. Even in so-called helical ferredoxins, the cysteine residues coordinating the cluster are coming from loops or beta-structured parts of the protein chain. In cases where one or two coordinating cysteines are in a helical part, it is usually the terminal part of the helix [15].

Until now, several successful FeS protein designs have been reported, but all of these relied on reproducing the natural fold around the bound FeS center. Coldren et al. incorporated a Fe_4S_4 binding site into the hydrophobic core of thioredoxin [16], whereas Nanda et al. [17], designed de novo a minimal rubredoxin fold. Other designs [18–

* Corresponding author. Tel.: +972 8 934 3505; fax: +972 8 934 4181.
E-mail address: dror.noy@weizmann.ac.il (D. Noy).
[1] Current address: Group of Biological Physics, Institute of Physics, Polish Academy of Science, Warsaw, Poland.

0005-2728/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbabio.2009.12.012



20] used sequences of natural FeS binding loops either isolated as short peptides, or grafted on top of a four-helix bundle scaffold preventing their assembly as multiple, adjacent FeS centers in electron-transfer chains.

Here, we introduce a new approach, computationally designing a highly regular and completely non-natural fold for a Fe_4S_4 cluster protein: a coiled-coil, single-chain four-helix bundle. Incorporating Fe_4S_4 binding sites into four-helix bundle proteins, the most robust and best characterized templates for de novo designed proteins, is ideal for constructing multi-center redox proteins. The fold repeats itself every seven residues with ~11 Å separation between repeating elements, hence, the parameterized, and symmetric backbone topology should enable extending the single binding site into a multi-center design by simply duplicating the same motif along the coiled-coil axis [21,22].

2. Materials and methods

2.1. Protein design

In this project, the design target was a four-helix bundle that coordinates a Fe_4S_4 cluster in the protein core. This was accomplished in multiple stages (Fig. 1). The first step was a survey of natural proteins to identify potentially useful motifs. Second, elements of symmetry in the cluster were identified that could constrain the conformation of the protein and simplify its design. Next, a library of possible backbones was generated by sampling a range of interhelical orientations and distances constrained by cofactor symmetry. In the fourth stage, this library was scanned for backbone and primary ligand conformations that were compatible with a cofactor-binding motif. Finally, loops were constructed and the sequence patterned using a collection of computational tools.

2.1.1. Step 1: Identify useful motifs

In a manual analysis of natural Fe_4S_4 proteins, we identified a minimal helical element that could be exploited in the design of a cofactor-binding helical bundle. We chose the Fe_4S_4 binding motif from *Thermatoga maritima* tryptophanyl-tRNA synthase (PDB ID 2G36) as a starting point because the cluster is located at the interface of three helical elements, and two of the four thiolate ligands were contributed by Cys 266 and Cys 269 of one α-helix (Fig. 1 — Step 1). A CXXC motif was utilized as a cluster binding half-site in the final design.

2.1.2. Step 2: Establish symmetry constraints

Structural analysis of metalloproteins has shown that the local symmetry of the metal-ligand coordination sphere is often related to the symmetry of the overall protein fold [23]. This approach has simplified design of other de novo metalloproteins [17,21,24]. Although the tryptophanyl-tRNA synthase does not show matched binding site and backbone symmetry, the atoms of the iron–sulfur cluster, and the sulfurs of the cysteine ligands form a cuboidal complex with tetrahedral symmetry (Fig. 1 — Step 2). This can be simplified to C2 symmetry along an axis formed by the midpoints of opposing edges of the tetrahedron. Two helices with CXXC motif could then fully complex the Fe_4S_4 cluster.

In order to match the protein topology to the cofactor symmetry, a library of ideal C2-symmetric tetrameric backbones was generated in protCAD [25] with four 3.5-residue/turn polyalanine helices aligned on the *z*-axis. Eight parameters defined the geometry of a C2 tetramer (four for each helix pair, Fig. 1 — Step 3): the bundle radius, the offset displacement between antiparallel helix pairs, the rotational phase of the helix, and the interhelical angle. These parameters were used to place one helix of each pair. The remaining two helices were created

Fig. 1. Computational design of a four-helix bundle coordinating a Fe_4S_4 cluster in its protein core.

by copying the first helix and rotating 180° around the *z*-axis. Once all the helices were placed, a coil transformation of pitch 190 Å was applied [26].

A library of backbones created by discretely sampling parameters was scanned for possible cofactor-binding capabilities (Fig. 1 — Step 4). One backbone ($r = 7.0$ Å, cluster binding $r' = 8.0$ Å, $\theta + \theta' = 100°$) that allowed for reasonable metal-ligand geometries (iron–sulfur bonds near 2.2 Å, C–S–Fe angles near 109.5°) was selected for further sequence patterning and loop design.

In order to create a single-chain protein, loops were introduced using protCAD (Fig. 1 — Step 5). Two canonical helix-turn-helix motifs were used: an α_L-β loop and a β-α_R-β loop [27,28]. Core and surface amino acids were patterned using a combination of protCAD for core amino acids and those packing against the ligand, and ROSETTA for surface exposed positions [29].

2.2. Protein preparation

DNA encoding the designed protein was optimized for *E. coli* expression system, and synthesized by Bio S&T (Canada). The gene was cloned into Pet32b vector (Trx-6xHis-tag, New England BioLabs) between KpnI and XhoI sites. Point mutations exchanging cysteine residues with serine (C15S, C19S, C63S, C67S) were performed on the gene template cloned into the pet32 vector. Mutagenesis was done in a one-step PCR reaction, using QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene). Point mutations were introduced by primers, designed as advised in the kit's manual. New DNA vectors were sequenced to confirm completed mutagenesis.

Proteins were expressed in the Bl21 *E. coli* strain, grown in LB medium + 100 μg/ml ampicilin, in 37 °C with vigorous shaking. When OD_{600} reached 0.6, IPTG was added to the culture to a final concentration of 0.2 mM. After three hours, the bacteria were harvested by centrifugation ($10\ \text{min} \times 6000 \times g$), and the resulting pellet was stored in −20 °C. The pellet was resuspended in buffer A (25 mM Tris/HCl pH 7.5, 0.5 M NaCl, 5 mM dithiotrietol [DTT]), enriched in 1 mM PMSF and 100 μg/l DNase, lysed by sonicating three times for 30 s, and incubated on ice for 30 min with occasional mixing. The lysate was centrifuged ($100{,}000 \times g \times 30$ min, 4 °C) and the supernatant was loaded on a HisTrap column (Amersham) that was previously equilibrated with buffer A. Protein fractions were eluted with a step gradient of 10 mM, 50 mM, and 300 mM imidazole in buffer A. SDS-PAGE electrophoresis verified that pure protein, was eluted in the final 300 mM imidazole step. This fraction was desalted (HiTrap Desalting, GE Healthcare) to 50 mM Tris/HCl, pH 7.5, and digested with TEV protease in order to cleave the Trx-6xHis-tag off the fused protein (1:30, TEV:fused protein in 50 mM Tris/HCl, 5 mM DTT, 0.01% TritonX-100 and 0.5 mM EDTA). The mixture was incubated at room temperature for 12 h with gentle stirring, and the purified cleaved protein was obtained from this mixture by a second HisTrap purification cycle. Before loading on the column, EDTA was neutralized by adding $MgCl_2$ to a final concentration of 2 mM. The purified protein was eluted during the initial gradient steps, leaving behind bound Trx-6xHis, TEV and small amounts of undigested fused protein. The flow-through fraction containing purified protein was concentrated with 5 kDa cut-off concentrators (VivaSpin), and its buffer was exchanged into 50 mM HEPES/NaOH, pH 8.0. Protein concentrations were determined spectrophotometrically, using an extinction coefficient of $\varepsilon_{280} = 16{,}500\ \text{M}^{-1}\,\text{cm}^{-1}$ calculated from the protein sequence. The purity and integrity of the protein were evaluated by Tricine-SDS-PAGE using 15% lower gel, 10% intermediate gel, 4% stacking gel, prepared and run as described in [30], and followed by Coomassie staining.

2.3. Mass spectrometry and in-gel digestion

Mass spectrometry (MS), in-gel tryptic digestion, and tandem MS analysis were performed in the Biological Mass Spectrometry facility at the Weizmann Institute of Science. Electroelution was performed in GeBAflex — tubes (Gene Bio Application Ltd., Israel) at 150 V for 2 h. Elution buffer contained 0.2–0.5% SDS, Tris and Tricine, pH 8.0. SDS removal after electroelution was performed by cold TCA:acetone precipitation in the presence of 1% sodium deoxycholate (NaDOC) [31]. Electrospray ionization (ESI) for intact molecular mass measurement was performed on an API 300 triple quadrupole mass spectrometer (ABI, PE Sciex, Concord, Ontario). Protein mass spectra were acquired using a nano-electrospray ion source (Protana A/S, Denmark). Before ESI-MS, the sample was passed over a micro-column consisting of about 300 nl Poros R1 reversed phase material (ABI). The purified protein was eluted directly into a nano-electrospray capillary (Proxion, Denmark) using 1 μl of 70% acetonitrile/5% formic acid mixture.

Protein bands excised from the SDS-PAGE were destained using multiple washings with 50% acetonitrile in 50 mM ammonium bicarbonate, subsequently reduced, alkylated, and in-gel digested with bovine trypsin (sequencing grade, Roche Diagnostics, Germany) at a concentration of 12.5 ng/μl in 50 mM ammonium bicarbonate at 37 °C. Peptide mixtures were extracted with 80% CH_3CN, 1% CF_3COOH, and the organic solvent was evaporated in a vacuum centrifuge. The resulting peptide mixtures were reconstituted in 80% Formic Acid and immediately diluted 1:10 with Milli-Q water. Peptide mixtures were separated by on-line reversed phase nanoscale capillary LC and analyzed by ESI-MS/MS. The samples were injected onto an in-house made 15 cm reversed phase spraying fused-silica capillary column (inner diameter 75 m, packed with 3 m ReproSil-Pur C18A18 media (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany), using an UltiMate 3000 Capillary/Nano LC System, consisting of FamosTM Micro Autosampler, SwitchosTM Micro Column Switching Module (LC Packings, Dionex). The LC setup was connected to the LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nano-electrospray ion source (Thermo Fisher Scientific, Bremen, Germany). The flow rate through the column was 250 nl/min. An acetonitrile gradient was employed with a mobile phase containing 0.1% and 0.2% formic acid in Milli-Q water in buffers A and B, respectively. The injection volume was 5 μl. The peptides were separated with 50 min gradients from 5 to 65% CH_3CN. In the nano-electrospray ionization source, the end of the capillary from the nano-LC column was connected to the emitter with pico-tip silica tubing, i.d. 20 μm (New Objective) and a stainless steel union, with a PEEK sleeve for coupling the nanospray with the on-line nano-LC. The voltage applied to the union in order to produce an electrospray was 2.4 kV. Helium was introduced as a collision gas at a pressure of 3 PSI. The mass spectrometer was operated in the data-dependent mode. Survey MS scans were acquired in the Orbitrap with the resolution set to a value of 60,000. Up to the seven of the most intense ions per scan were fragmented and analyzed in the linear trap. For the analysis of tryptic peptides, survey scans were recorded in the FT-mode followed by data-dependent collision-induced dissociation (CID) of the seven most intense ions in the linear ion trap (LTQ). Raw spectra were processed using open-source software DTASuperCharge (http://msquant.sourceforge.net). The data were searched with MASCOT (Matrix Science, London, UK) against a Swissprot or NCBI database + the target protein sequence. Search parameters included variable modifications of 57.02146 Da (carboxyamidomethylation) on Cys, 15.99491 Da (oxidation) in Met and 0.984016 Da (deamidation) on Asn and Gln. The search parameters were as follows: maximum 2 missed cleavages, initial precursor ion mass tolerance 10 ppm and fragment ion mass tolerance 0.6 Da. The identity of the peptides was deduced from the detected collision-induced dissociation products by the Mascot and Sequest programs and confirmed by manual inspection of the fragmentation series.

2.4. Protein assembly with FeS clusters

Iron–sulfur clusters were incorporated into the apoprotein by a modified procedure based on Scott, et al. [20,32]. A solution of 200 μM

apoprotein in 50 mM HEPES/NaOH, pH 8.0 was deoxygenated by four cycles of vacuum pumping/nitrogen filling, and about an hour incubation with stirring under anaerobic atmosphere. The deaerated solution was placed in an anaerobic chamber and all the following steps were performed in strictly anaerobic atmosphere, at ambient temperature. DTT was added from a 50 mM water stock to a final concentration of 2 mM (DTT:apoprotein ratio 10:1) and incubated for one hour. Then, $FeCl_3$ and Na_2S were sequentially added from their 50 mM water stocks to a final concentration of 1 mM (reactant: apoprotein ratio 5:1), and incubated for about 15 min. Longer incubation did not yield higher reaction efficiency. Excess DTT, $FeCl_3$ and Na_2S were removed by passing the red-brown reaction mixture through a sephadex G-10 (GE Healthcare) column, equilibrated with 50 mM HEPES/NaOH, pH 8.0.

2.5. Elemental analysis

Holoprotein samples from three different preparations were each split in two aliquots. The first batch of aliquots containing 0.4–1 ml of sample solutions was digested for 1 h in 2 ml of 65% HNO_3 at 90 °C, and then incubated overnight at 70 °C in this acidic mixture. Vessels were then cooled down and water was added to a total volume of 10 ml. Iron concentrations in these clear solutions were measured using an End-On-Plasma ICP/AES model 'ARCOS' (Spectro GMBH, Germany). Measurements were calibrated with ICP standards (Merck). The second batch of aliquots was used for determining protein concentration by the fluorescamine (Sigma)-based fluorometric assay. The protein samples were diluted to concentrations of 0.5–5 μM in 50 mM HEPES pH 8, and incubated with 500 μM fluorescamine (added from 20 mM acetone stock) for 60 min at room temperature. An apoprotein sample with known concentration was prepared in the same way for using as a calibration standard. Samples were excited at 390 nm and an emission spectrum was measured between 400 to 600 nm [33]. The concentration was calculated from the fluorescence peak at 490 nm using a calibration curve obtained from serial dilution of the apoprotein sample.

2.6. Gel filtration

The sizes of the apo- and holoprotein were determined by gel filtration using a FPLC system (Akta Purifier, GE healthcare) and a Sephadex75 column. The column was equilibrated with Tris/HCl pH 7.5 + 0.5 M/1 M NaCl buffer solutions, with or without 1 mM DTT. Holoprotein samples were analyzed in buffers without DTT, to prevent loss of iron–sulfur cluster. Anaerobic conditions were maintained by bubbling the buffer with nitrogen, and pre-equilibrating the column with the deaerated buffer for at least two column volumes. Samples were loaded through a 1 ml loop. Chromatograms were recorded by monitoring the absorbance simultaneously at 280 nm and 415 nm. The molecular mass was calculated from a calibration curve made by using a low molecular weight (LMW) Calibration Kit (GE Healthcare).

2.7. Analytical ultracentrifugation

Sedimentation velocity experiments were carried out in Beckman's XLA analytical ultracentrifuge equipped with UV–Vis absorbance optics. Standard two-channel cells (spin analytical, Austin TX) were assembled, filled with water and tested for leaks by spinning at 50,000 rpm. Then, water was taken out; the cells were filled with sample solution in an anaerobic chamber and sealed with the standard polypropylene gaskets and brass screws. The cells were immediately put in the centrifuge rotor and placed in the ultracentrifuge under vacuum. Sedimentation velocity measurements started after at least two hours of equilibration at 20 °C with the rotor at rest. All measurements were carried out at 50,000 rpm, detection wavelengths were 290 nm for apoprotein samples and 420 nm for holoprotein samples. Sedimentation coefficient distributions, $c(s)$, were obtained from the raw data by using the software Sedfit [34] (obtained from http://www.analyticalultracentrifugation.com) using a theoretical partial specific volume of 0.7223 ml g^{-1}, calculated from its sequence using the amino acid residue parameters of Kharakoz [35] substituted for the Cohn and Edsall parameters [36] in the program "Sednterp" [37]. The latter program was also used for calculating solvent densities and viscosities from solvent composition.

2.8. Optical spectroscopy

UV–Vis absorption and CD spectra were recorded by JASCO V-7200, and J-810 spectrometers, respectively, in gas tight cuvettes (10 mm optical pathlength, 1 mm for UV-CD).

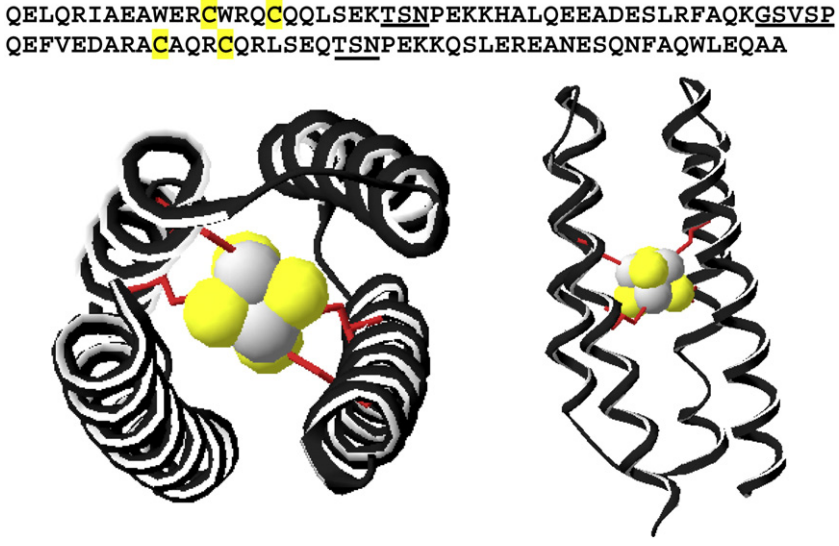


Fig. 2. Sequence, and modeled structure of CCIS1, top and side views. Loop residues are underlined. Cysteines are highlighted in the sequence and represented as red sticks in the structure. The Fe_4S_4 cluster atoms are shown as spheres (iron in white, sulfur in yellow).

2.9. EPR spectroscopy

Low temperature X band EPR spectra were measured on a Bruker ELEXSYS 500 spectrometer in 3.4 mm (i.d.) capillary tubes. Temperature was controlled by a cryostat cooled by liquid helium (Oxford Instruments). EPR signals from 2,2-diphenyl-1-picrylhydrazyl (DPPH) were used as reference for determining *g* values.

3. Results

3.1. Protein design and preparation

The 'metal-first' approach for computational design of metalloproteins [17,24], provided CCIS1 (*C*oiled-Coil *I*ron *S*ulfur 1), an all α-helical protein scaffold in which a single Fe_4S_4 cluster is embedded within a stable four-helix bundle core (Fig. 2). Placing iron-coordinating cysteines at adjacent *d* and *a* positions within the hydrophobic coiled-coil core makes this new protein different from any known natural FeS protein fold, and exceptional among previous de novo designs [17–20] that relied on natural folds and sequences.

CCIS1 apoprotein was prepared and purified by standard recombinant DNA and protein preparation techniques. Overexpression and purification of a synthetic gene encoding the designed protein sequence yielded a water-soluble protein. Tricine-SDS-PAGE analysis revealed two bands of apparent mass close to 11 kDa. The minor band was less than 25% of the total protein, as calculated by gel densitometry [38]. MS confirmed that the major band has the expected molecular weight of CCIS1, namely 11.661 kDa, and the minor band has a molecular weight of 9.557 kDa; in-gel-tryptic digestion and tandem MS analysis indicated that both bands were intact, and truncated CCIS1, respectively. The identified peptides coverage was 61%, and included both the C-terminal and a part close to the N-terminal of the protein (Scheme 1). MS of Trx-CCIS1 revealed similar sample composition: a major fraction with the expected mass of Trx-CCIS1, and a minor fraction, about 2 kDa smaller. This indicates that truncation occurs already at the expression stage.

3.2. Protein folding and assembly with Fe_4S_4 clusters

A Fe_4S_4 cluster was incorporated into the apoprotein under anaerobic conditions by *in situ* chemical synthesis according to published protocols [20,32]. The yield of reconstitution was about 75%. A variety of spectroscopic and analytical methods was employed in order to characterize the FeS center, and verify that the protein folds are assembled as designed. Importantly, comparing the UV-circular dichroism (UV-CD) spectra of the apo- and holoprotein clearly indicates that the latter is mostly α-helical, and assembly with the Fe_4S_4 cluster significantly stabilizes this fold (Fig. 3). This implies that the α-helical fold is required for proper coordination of the Fe_4S_4 cluster by the ligating cysteines.

The absorption spectrum of the holoprotein (Fig. 4) has the typical features of Fe_4S_4 cluster proteins; namely, a peak at 415 nm with a shoulder at 360 nm, which disappear upon reduction with dithionite. Under the same conditions for assembling Fe_4S_4 CCIS1-4S, a CCIS1 mutant, in which all four cysteine residues are replaced by serine, has no indication for specific binding (Fig. 4). Gel filtration and sedimentation velocity measurements (Fig. 5) confirmed that the 415 nm absorbing species are protein complexes. Additionally, the CD

1 QELQR**IAEAW ERCWR**QCQQL SEKTSNPEKK **HALQEEADES LRFAQKGSVS**

51 **PQEFVEDARA CAQR**CQRLSE QTSNPEK**KQS LEREANESQN FAQWLEQAA**

Scheme 1. CCIS1 sequence and coverage by MS/MS identified peptides. Matched peptides are underlined.

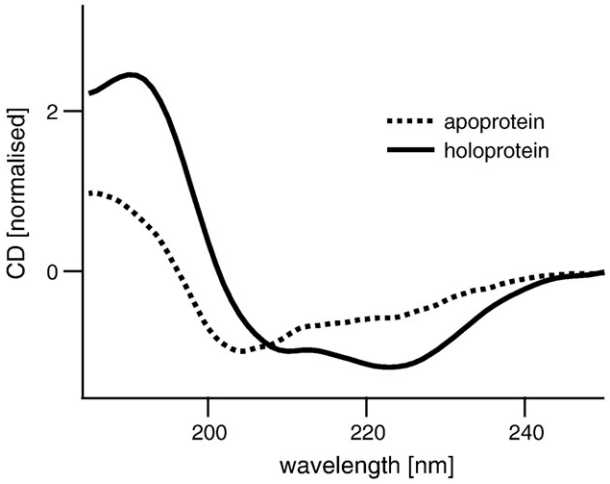


Fig. 3. Comparison of UV-CD spectra of apo- and holo-CCIS1 reveals increased α-helical nature upon assembly with a Fe_4S_4 cluster.

spectrum (Fig. 4 inset) in the visible wavelength range has typical ferredoxin features with a narrow negative signal centered at 360 nm, and a broader one centered at 560 nm. We determined an extinction coefficient of 18,000 M^{-1} cm^{-1} at 415 nm for the holoprotein, which is within the range of 16,000 to 23,000 M^{-1} cm^{-1} reported for natural proteins incorporating one Fe_4S_4 cluster [39]. Furthermore, elemental analysis of the holoprotein indicates a ratio of 4.2 ± 0.4 Fe atoms per protein. These results strongly suggest complete occupation of all Fe_4S_4 binding sites in holo-CCIS1.

Gel filtration and sedimentation velocity measurements indicate a significant fraction of dimers and higher oligomers in apo- and holoprotein samples (Fig. 5). However, addition of DTT to the apoprotein almost completely eliminated the non-monomeric species (Fig. 5A, C). Chromatograms and *c*(s) distributions of holoprotein

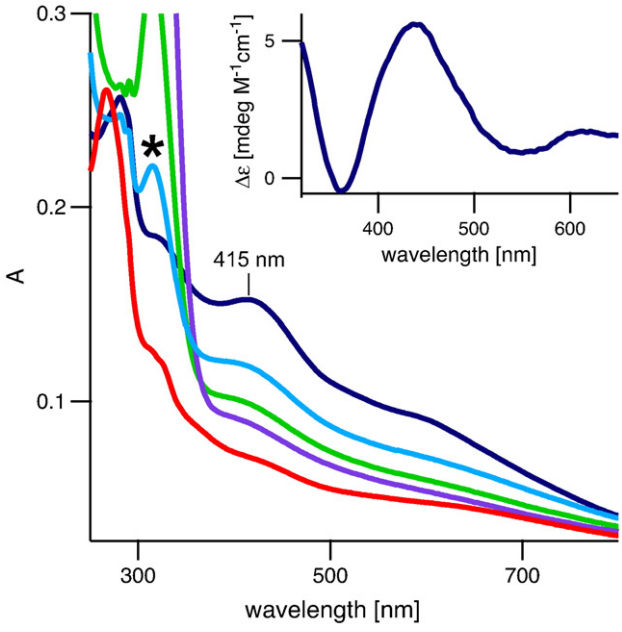


Fig. 4. Visible absorption (main) and CD (inset) spectra of 10 μM holo-CCIS1 prior to reduction with dithionite (dark blue). Reduction leads to disappearance of the 415 nm peak (cyan, green, and purple lines, the asterisk marks the dithionite absorbance peak). The spectrum of the CCIS-4S mutant (red), indicates no specific Fe_4S_4 cluster binding.

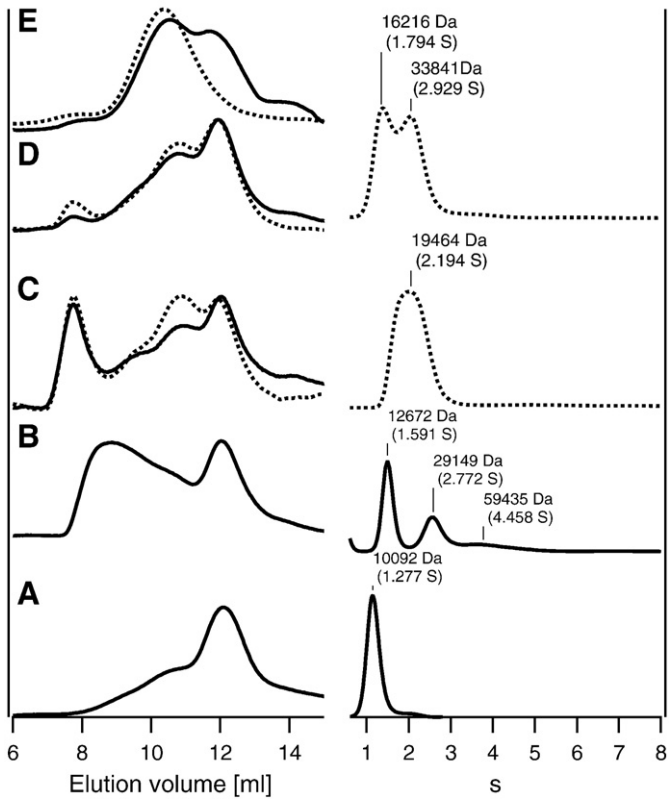


Fig. 5. Typical gel filtration (left) and sedimentation coefficient distributions (right), $c(s)$, of apo-CCIS1 in the absence (A) and presence (B) of DTT, holo-CCIS1 in the absence (C) and presence (D) of 1 M NaCl, and holo-CCIS1-2S in 1 M NaCl. Dashed, and solid lines indicate chromatograms and sedimentation curves that were recorded by monitoring absorbance at 280 nm, and 415 nm, corresponding to absorption peaks of protein residues, and the Fe_4S_4 cofactor, respectively.

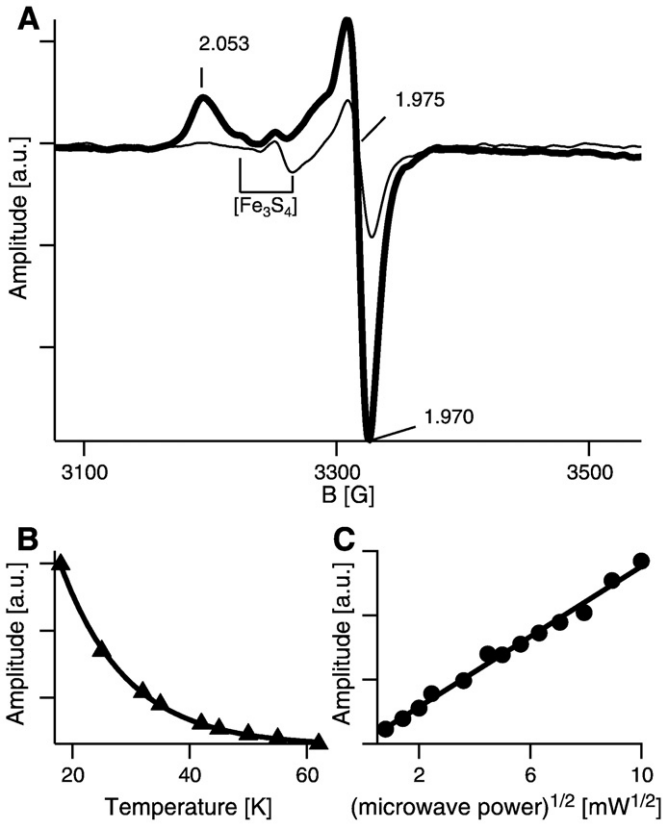


Fig. 6. A) EPR spectra of 30 μM holo-CCIS1 before (thick line) and after (thin line) reduction with dithionite, recorded at $T = 20$ K. B) Temperature dependence of the EPR signal amplitude (peak-to-peak) from the reduced sample, measured at a microwave power of 15 mW, and modulation amplitude of 1 G. C) The respective power dependence at $T = 20$ K.

samples (Fig. 5B, C) revealed the presence of oligomers, dimers and monomers. However, the characteristic Fe_4S_4 cluster absorbance of 415 nm is detected primarily in fractions corresponding to dimers and monomers. Furthermore, the monomeric fraction becomes predominant by assembling the protein in the presence of 1 M NaCl.

Interestingly, the apoprotein monomer-dimer distribution of the CCIS1-C15S-C19S mutant (CCIS1-2S) in which two cysteines were replaced by serines, is similar to CCIS1. However, in the holoprotein of this mutant the characteristic Fe_4S_4 cluster absorbance of 415 nm is detected exclusively in fractions corresponding to dimers (Fig. 5B). Formation of a CCIS1-2S dimer bound to Fe_4S_4 clusters indicates that one or several potential domain-swapped configurations are being adopted. A strategy for eliminating dimer formation by stabilizing turn regions or by redesigning helix–helix interactions to enhance specificity [40,41], will be outlined in the discussion section.

3.3. Characteristics of the CCIS1-bound Fe_4S_4 cluster

The properties of the FeS center in holo-CCIS1 were characterized by EPR spectroscopy (Fig. 6). Dithionite-reduced samples show the expected characteristic rhombic spectra, typical of low-potential $[Fe_4S_4]^+$ ($S = 1/2$), with g values of $g_{z,y} = (1.970, 1.975)$, and $g_x = 2.053$ [7]. The temperature dependence of the signal intensity (Fig. 6B) was also characteristic of $[Fe_4S_4]^+$ clusters with the signal almost vanishing at $T > 50$ K. The signal's linear power dependence (Fig. 6C) indicated that spin relaxation is very efficient, but there is hardly any interaction between adjacent clusters [42].

Notably, a weak EPR signal (about 30% of the reduced sample's amplitude) is detected without dithionite reduction. This may suggest that the complex is partially reduced by dithiothreitol (DTT) already during the assembly stage. An alternative explanation is the formation of a $[Fe_3S_4]^+$ sub-fraction as a result of oxidative damage leading to the loss of one Fe(II) atoms [43]. However, this should give rise to a distinct EPR spectrum [44,45]. In our case, most of the signal in unreduced samples is typical of $[Fe_4S_4]^+$ except for a minor fraction around $g = 2.03$. The presence of the $S = 1/2$ state of Fe_4S_4 under these conditions in natural proteins is very unusual but has been observed in thioredoxin reductase [46]. In that case a thyil radical of a fifth cysteine in close proximity to the cluster gives rise to the EPR signal. In our design, a fifth cysteine is not available and the origin of the EPR signal need to be further explored. A possible explanation is formation of thyil radicals by DTT that is incorporated near the Fe_4S_4 binding site during the holoprotein assembly process.

Apparently, the $[Fe_4S_4]^+$ state of the new protein is unstable and reduction is irreversible. This is probably due to hydrolysis of the sulfur atoms [47] upon reduction of the cluster. We are planning to improve stability in the next design by adding specific second shell hydrogen bonds to the sulfur atoms from nearby protein residues [18]. The stability of the rubredoxin mimic, RM1, to multiple redox cycles was attributed to a network of backbone amide to cysteine thiol hydrogen bonds [17].

4. Discussion

Altogether, results from a wide range of analytical and spectroscopic measurements verify that the protein is folded and assembled as designed, namely, an α-helical coiled-coil binding a single Fe_4S_4 cluster. The new design is very promising although optimization is required, particularly in preventing domain-swapped dimers, and

stabilizing the $[Fe_4S_4]^+$ state. The atomic model of the current design (Fig. 2) allows rational, structure-based engineering of the next design. Possible strategies to consider include (1) redesigning surface electrostatics to enhance specificity of helix–helix interactions, (2) improving turn stability to prevent domain-swapped conformers [48], and (3) improving helix stability by adding helix-favoring residues, thus enhancing binding site preorganization and $[Fe_4S_4]^+$ stability.

Modifying the distribution of charged residues on the surface of CCIS1 may address issues such as aggregation. Modeling the electric field distribution shows discrete regions of positive and negative charge which could drive intermolecular associations (Fig. 7A). Additionally, the observation of dimers in the Cys-Ser mutants suggests that domain-swapped structures may be formed. Potential configurations are shown in Fig. 7B. One strategy for preventing alternate configurations is *negative design*. Charge pair interactions between helices will be chosen that stabilize helix–helix interfaces in the desired conformation while destabilizing 'non-native' helix–helix interfaces in competing configurations. These will be implemented in the next generation of iron–sulfur cluster designs.

The main advantage of the new design is its extendibility into a multi-FeS cluster protein by simply duplicating and translating the binding site along the coiled-coil axis. This opens new possibilities for constructing functional analogs of the natural multi-FeS redox chains that are found for example in hydrogenase and complex I. These may be used as simple models for studying the details of electron flow-through redox chains to catalytic sites. Additionally, they can be used in more practical applications as molecular "wires" for electron transfer between the active sites of any given set of redox enzymes.

5. Concluding remarks

In conclusion, we presented here the first computational de novo design of a Fe_4S_4 cluster protein with a non-natural α-helical fold. This was accomplished by novel computational methods that concurrently

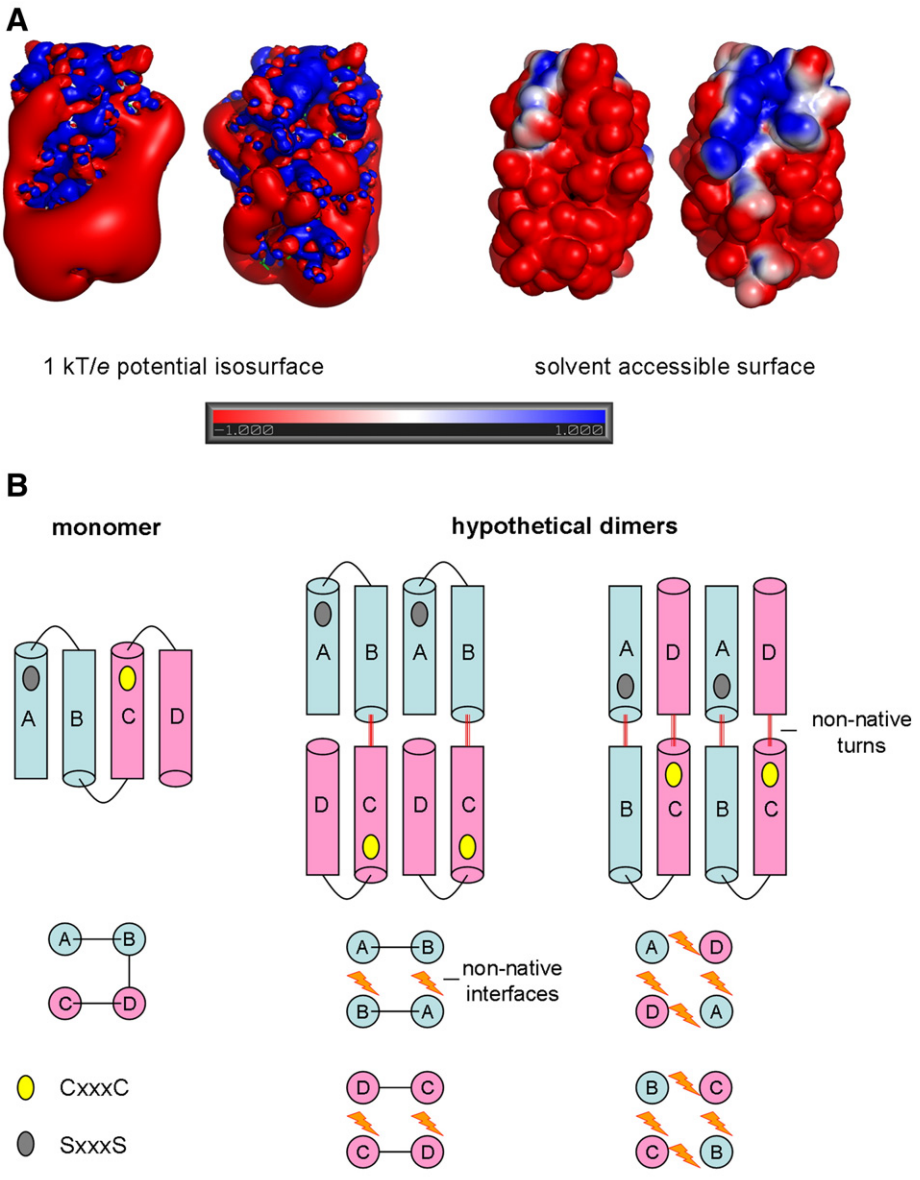


Fig. 7. A) Electrostatic potential surfaces of CCIS1 calculated using the Adaptive Poisson–Boltzmann Solver (APBS) [49]. B) Monomeric vs. possible dimeric configurations for CCIS1.

sampled backbone, sequence and sidechain space for an entire protein. The identification of both a sequence motif and a three-dimensional motif consistent with full coordination of a Fe_4S_4 cluster is a radical departure from all previous de novo designs. These consisted of either grafting loops from natural proteins or using unstructured, glycine-rich scaffolds, where binding is coupled with significant conformational restrictions. Thus, unlike many other designs which employ a trial-and-error protein engineering approach, here we have a clear, structural hypothesis upon which the design is based. Several key experiments support the central design hypothesis that a FeS cluster was successfully introduced into the hydrophobic core of a coiled-coil. The CD spectrum shows a clear transition from random-coil to α-helix upon addition of the iron–sulfur cofactor; mutation of one half-site to non-interacting residues resulted in an obligate dimer for cofactor binding; and EPR measurements indicate that a $[Fe_4S_4]^+$ cofactor is bound to the protein. Finally, our structure-based hypothesis and the atomic model of the design make it possible for us to conduct detailed post-hoc evaluation, thereby setting forth a number of practical strategies for improving the shortcomings of the first prototype in the next generation of designs.

Acknowledgements

Dror Noy acknowledges support from the Human Frontiers Science Program Organization. Vikas Nanda and Fei Xu wish to recognize support from the NSF (DMR-0907273).

References

[1] J. Meyer, Iron–sulfur protein folds, iron–sulfur chemistry, and evolution, J. Biol. Inorg. Chem. 13 (2008) 157–170.
[2] G. Shen, J.H. Golbeck, Assembly of the bound iron–sulfur clusters in photosystem I, in: J.H. Golbeck (Ed.), Photosystem I: the Light-Driven Plastocyanin:Ferredoxin Oxidoreductase, Springer, 2006, pp. 529–547.
[3] T. Friedrich, K. Steinmüller, H. Weiss, The proton-pumping respiratory complex I of bacteria and mitochondria and its homologue in chloroplasts, FEBS Lett. 367 (1995) 107–111.
[4] M.L. Ghirardi, M.C. Posewitz, P.C. Maness, A. Dubini, J. Yu, M. Seibert, Hydrogenases and hydrogen photoproduction in oxygenic photosynthetic organisms, Annu. Rev. Plant Biol. 58 (2007) 71–91.
[5] J.C. Fontecilla-Camps, A. Volbeda, C. Cavazza, Y. Nicolet, Structure/function relationships of [NiFe]- and [FeFe]-hydrogenases, Chem. Rev. 107 (2007) 4273–4303.
[6] W. Lubitz, E. Reijerse, M. van Gastel, [NiFe] and [FeFe] hydrogenases studied by advanced magnetic resonance techniques, Chem. Rev. 107 (2007) 4331–4365.
[7] M.S. Koay, M.L. Antonkine, W. Gartner, W. Lubitz, Modelling low-potential [Fe4S4] clusters in proteins, Chem. Biodivers. 5 (2008) 1571–1587.
[8] W. Lubitz, E.J. Reijerse, J. Messinger, Solar water-splitting into H_2 and O_2: design principles of photosystem II and hydrogenases, Energy Environ. Sci. 1 (2008) 15–31.
[9] R.A. Grimme, C.E. Lubner, D.A. Bryant, J.H. Golbeck, Photosystem I/molecular wire/metal nanoparticle bioconjugates for the photocatalytic production of H-2, J. Am. Chem. Soc. 130 (2008) 6308–6309.
[10] M. Ihara, H. Nishihara, K.S. Yoon, O. Lenz, B. Friedrich, H. Nakamoto, K. Kojima, D. Honma, T. Kamachi, I. Okura, Light-driven hydrogen production by a hybrid complex of a [NiFe]-hydrogenase and the cyanobacterial photosystem I, Photochem. Photobiol. 82 (2006) 676–682.
[11] M. Ihara, H. Nakamoto, T. Kamachi, I. Okura, M. Maeda, Photoinduced hydrogen production by direct electron transfer form photosystem I cross-linked with cytochrome *c*3 to [NiFe]-hydrogenase, Photochem. Photobiol. 82 (2006) 1677–1685.
[12] R.L. Koder, K.G. Valentine, J. Cerda, D. Noy, K.M. Smith, A.J. Wand, P.L. Dutton, Nativelike structure in designed four alpha-helix bundles driven by buried polar interactions, J. Am. Chem. Soc. 128 (2006) 14450–14451.
[13] B. Kuhlman, G. Dantas, G.C. Ireton, G. Varani, B.L. Stoddard, D. Baker, Design of a novel globular protein fold with atomic-level accuracy, Science 302 (2003) 1364–1368.
[14] A.M. Burroughs, S. Balaji, L.M. Iyer, L. Aravind, Small but versatile: the extraordinary functional and structural diversity of the β-grasp fold he extraordinary functional and structural diversity of the β-grasp fold, Biol. Direct 2 (2007) 18.
[15] S.S. Krishna, R.I. Sadreyev, N.V. Grishin, A tale of two ferredoxins: sequence similarity and structural differences, Bmc Struct. Biol. 6 (2006).
[16] C.D. Coldren, H.W. Hellinga, J.P. Caradonna, The rational design and construction of a cuboidal iron–sulfur protein, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 136635–136640.
[17] V. Nanda, M.M. Rosenblatt, A. Osyczka, H. Kono, Z. Getahun, P.L. Dutton, J.G. Saven, W.F. DeGrado, De novo design of a redox-active minimal rubredoxin mimic, J. Am. Chem. Soc. 127 (2005) 5804–5805.
[18] M.L. Antonkine, M.S. Koay, B. Epel, C. Breitenstein, O. Gopta, W. Gärtner, E. Bill, W. Lubitz, Synthesis and characterization of de novo designed peptides modelling the binding sites of [4Fe–4S] clusters in photosystem I, Biochim. Biophys. Acta, Bioenerg. 1787 (2009) 995–1008.
[19] B.R. Gibney, S.E. Mulholland, F. Rabanal, P.L. Dutton, Ferredoxin and ferredoxin-heme maquettes, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 15041–15046.
[20] M.P. Scott, J. Biggins, Introduction of a [4Fe–4S (S-cys)4](+1,+2) iron–sulfur center into a four-alpha helix protein using design parameters from the domain of the F(x) cluster in the Photosystem I reaction center, Protein Sci. 6 (1997) 340–346.
[21] F.V. Cochran, S.P. Wu, W. Wang, V. Nanda, J.G. Saven, M.J. Therien, W.F. DeGrado, Computational de novo design and characterization of a four-helix bundle protein that selectively binds a nonbiological cofactor, J. Am. Chem. Soc. 127 (2005) 1346–1347.
[22] K.A. McAllister, H. Zou, F.V. Cochran, G.M. Bender, A. Senes, H.C. Fry, V. Nanda, P.A. Keenan, J.D. Lear, J.G. Saven, M.J. Therien, J.K. Blasie, W.F. DeGrado, Using alpha-helical coiled-coils to design nanostructured metalloporphyrin arrays, J. Am. Chem. Soc. 130 (2008) 11921–11927.
[23] A. Lombardi, C.M. Summa, S. Geremia, L. Randaccio, V. Pavone, W.F. DeGrado, Inaugural article: retrostructural analysis of metalloproteins: application to the design of a minimal model for diiron proteins, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 6298–6305.
[24] C.M. Summa, A. Lombardi, M. Lewis, W.F. DeGrado, Tertiary templates for the design of diiron proteins, Curr. Opin. Struct. Biol. 9 (1999) 500–508.
[25] C.M. Summa, Computational methods and their applications for de novo functional protein design and membrane protein solubilization, Doctoral Thesis (2002) University of Pennsylvania School of Medicine, Philadelphia.
[26] S.F. Betz, W.F. DeGrado, Controlling topology and native-like behavior of de novo-designed peptides: design and characterization of antiparallel four-stranded coiled coils, Biochemistry 35 (1996) 6955–6962.
[27] S.J. Lahr, D.E. Engel, S.E. Stayrook, O. Maglio, B. North, S. Geremia, A. Lombardi, W.F. DeGrado, Analysis and design of turns in alpha-helical hairpins, J. Mol. Biol. 346 (2005) 1441–1454.
[28] A.V. Effimov, Patterns of loop regions in proteins, Curr. Opin. Struct. Biol. 3 (1993) 379–385.
[29] B. Kuhlman, D. Baker, Native proteins sequences are close to optimal for their structures, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 10383–10388.
[30] H. Schagger, Tricine-SDS-PAGE, Nat. Protoc. 1 (2006) 16–22.
[31] C. Montigny, C. Jaxel, A. Shainskaya, J. Vinh, V. Labas, J.V. Moller, S.J.D. Karlish, M. le Maire, Fe2+-catalyzed oxidative cleavages of Ca2+-ATPase reveal novel features of its pumping mechanism, J. Biol. Chem. 279 (2004) 43971–43981.
[32] T.C. Sow, M.V. Pedersen, H.E.M. Christensen, B.L. Ooi, Total synthesis of a miniferredoxin, Biochem. Biophys. Res. Commun. 223 (1996) 360–364.
[33] S. Undenfriend, S. Stein, P. Bohlen, W. Dairman, Fluorescamine: a reagent for assay of amino acids, peptides, proteins and primary amines in the picomole range, Nature 178 (1972) 871–872.
[34] P. Schuck, Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling, Biophys. J. 78 (2000) 1606–1619.
[35] D.P. Kharakoz, Partial volumes and compressibilities of extended polypeptide chains in aqueous solution: additivity scheme and implication of protein unfolding at normal and high pressure, Biochemistry 36 (1997) 10276–10285.
[36] E.J. Cohn, J.T. Edsall, Proteins, Amino Acids, and Peptides, Reinhold, New York, 1943.
[37] T. Laue, B.D. Shaw, T.M. Ridgeway, S.L. Pelletier, Computer-aided interpretation of analytical sedimentation data for proteins, in: S.E. Harding, A.J. Rowe, J.C. Horton (Eds.), Analytical Ultracentrifugation in Biochemistry and Polymer Science, The Royal Society of Chemistry, Cambridge, U. K., 1992, pp. 90–125.
[38] M.D. Abramoff, P.J. Magelhaes, S.J. Ram, Image processing with ImageJ, Biophoton. Int. 11 (2004) 36–42.
[39] W.V. Sweeney, J.C. Rabinowitz, Proteins containing 4Fe–4S clusters: an overview, Annu. Rev. Biochem. 49 (1980) 139.
[40] J.J. Havranek, P.B. Harbury, Automated design of specificity in molecular recognition, Nat. Struct. Mol. Biol. 10 (2003) 45–52.
[41] C.M. Summa, M.M. Rosenblatt, J.K. Hong, J.D. Lear, W.F. DeGrado, Computational de novo design, and characterization of an A(2)B(2) diiron protein, J. Mol. Biol. 321 (2002) 923–938.
[42] N.J. Malmberg, J.J. Falke, Use of EPR power saturation to analyze the membrane-docking geometries of peripheral proteins: applications to C2 Domains, Annu. Rev. Biophys. Biomol. Struct. 34 (2005) 71–90.
[43] F. Capozzi, S. Ciurli, C. Luchinat, Coordination sphere versus protein environment as determinants of electronic and functional properties of iron sulphur proteins, Struct. Bond. 90 (1998) 128–160.
[44] R. Camba, F.A. Armstrong, Investigations of the oxidative disassembly of Fe–S clusters in Clostridium pasteurianum 8Fe ferredoxin using pulsed-protein-film voltammetry, Biochemistry 39 (2000) 10587–10598.
[45] J.A. Cowan, S.M. Lui, Structure–function correlations in high-potential iron proteins, Adv. Inorg. Chem. 45 (1998) 313–350.
[46] E.M. Walters, M.K. Johnson, Ferredoxin:thioredoxin reductase: disulfide reduction catalyzed via novel site-specific [4Fe–4S] cluster chemistry, Photosynth. Res. 79 (2004) 249–264.
[47] M.L. Kennedy, B.R. Gibney, Proton coupling to $[4Fe4S]^{2+/+}$ oxidation and reduction in a designed protein, J. Am. Chem. Soc. 12 (2002) 6826–6827.
[48] N.L. Ogihara, G. Ghirlanda, J.W. Bryson, M. Gingery, W.F. DeGrado, D. Eisenberg, Design of three-dimensional domain-swapped dimers and fibrous oligomers, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 1404–1409.
[49] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 10037–10041.